STATISTICAL DECISION
IN
LUNAR SITE CERTIFICATION

December 31, 1963

Bellcomm, Inc.
Washington, D.C.

by M.A. Leibowitz

TABLE OF CONTENTS

PURPOSE AND SCOPE

ABSTRACT

1.0 INTRODUCTION

2.0 BASIC IDEAS OF STATISTICAL DECISION THEORY

3.0 LUNAR SITE CERTIFICATION PROBLEM

4.0 SUMMARY

5.0 ACKNOWLEDGEMENTS

PURPOSE AND SCOPE

In determining the suitability of a given area on the moon for a lunar landing, prior observations may be of decisive importance. The purpose of this paper is to discuss the statistical decision problems arising in certifying a lunar area for landing on the basis of these observations. Criteria for evaluating any decision procedure are introduced, and a discussion is given of how these criteria may be satisfied.

ABSTRACT

This paper discusses the statistical decision problem arising in determining on the basis of prior observations, the suitability of a given lunar area for landing.

For a decision criterion we will use the notions of "catastrophic error" and "false alarm." In the present context, these refer, respectively, to a decision leading to disaster (i.e., a failure on landing) and to a rejection of an actually suitable surface.

A statistical method will be described which insures that the probability of a catastrophic error due to the observed properties is less than any prescribed value no matter what the lunar area is actually like, and it will be shown how the false alarm probability may be obtained. Finally, two examples will be analyzed in detail.

STATISTICAL DECISION IN LUNAR SITE CERTIFICATION

1.0 INTRODUCTION

In determining the suitability of a given area on the moon for a lunar landing any prior observations may be of decisive importance. Almost inevitably, such observations are to some extent, uncertain or incomplete. They may, for example be of limited parts of the area, be restricted to certain properties, or distorted by measurement errors. Decision making in the face of imperfect knowledge usually involves probabilistic or statistical considerations. This paper discusses the statistical decision problem arising in certifying for landing a lunar area.

For a decision criterion we will use the notions of "catastrophic error" and "false alarm". In the present context, these refer respectively to a decision leading to disaster (i.e., a failure on landing) and to a rejection of an actually suitable surface.

A statistical method will be described which insures that the probability of a catastrophic error due to the observed properties is less than any prescribed value no matter what the lunar area is actually like. It will also be shown how the false alarm probability may be obtained. Finally, two examples will be analyzed in detail.

2.0 BASIC IDEAS OF STATISTICAL DECISION THEORY

Before considering the lunar certification problem, we review briefly some of the ideas of statistical decision theory.

Suppose that we are given some statement or hypothesis whose truth or falsity is to be decided. Characteristic of the statistical decision theory is that the hypotheses considered are all to some extent of a statistical nature. Thus an hypothesis might be that a given coin is heavily biased toward the head side, i.e., probability of a head $>> 1/2$; or it might be that a group of electrical parts has a mean life greater than a given period; or it might be that a certain person is suffering from a certain disease. In this last instance, statistics may enter in the diagnosis; for example, the distribution of some component of the blood among normal people may be very different than the distribution among those with the disease.

In any case, in making a decision there are two different types of errors which can be made, namely:

(1) An hypothesis may be called false when it is actually true.

(2) An hypothesis may be called true when it is actually false.

The consequences of making an error of either type are generally not the same. Thus for the hypothesis above, that a person is suffering from a disease, an error of the first type (saying the patient is free of the disease when he isn't) may be fatal since it can mean that proper treatment is not instituted. On the other hand, an error of the second type (saying that the patient has the disease when he actually doesn't) may cause unnecessary alarm but is unlikely to have permanent effect.

Rather than use the undescriptive names, errors of Type I and Type II (customary in statistic books), it is more appropriate to use here the terms "catastrophic error" and "false alarm." To distinguish between the two, one may use the following rule: to ignore danger when present may be fatal; to proclaim danger when absent can be, at worst, highly alarming.

Since, obviously, a catastrophic error must be avoided at all reasonable cost, it is natural to impose on any decision procedure the requirement that the probability of making such an error be less than some very small number, say .01 or .05. Meeting this requirement by itself never imposes any problem. If a catastrophic error can result when a true hypothesis is called false, one can always call the hypothesis true (the patient always has the disease). A catastrophic error will never be made, but the "false alarm" rate may be so high that such an ultra-conservative attitude is completely impractical. The skill of the decision maker enters in meeting the requirement on the catastrophic error and yet keeping the false alarm rate low.

To put these ideas in a more quantitative form let H denote some hypothesis. The decision on whether or not H is true is made in the following way: an observation or measurement of some kind is performed; call the result of this measurement y. If y has <u>any one</u> of a certain set of values C, the hypothesis is concluded to be true; otherwise the hypothesis is called false. The set of values C is generally referred

to as the "critical region" since the truth or falsity of the hypothesis is decided on the basis of whether or not y is part of C.

The probability of y being in C (i.e., having one of the values included in C) depends on whether H is true or not. If H is true, let us denote the probability of y being in C by $P(C|H)$; and if H is false, by $P(C|\bar{H})$. Then:

I. The probability of calling a true hypothesis false is $1 - P(C|H)$.

In fact, by the definition of C, any hypothesis would be called false only if y was found to be outside of C; but this is one minus the probability of y being in C.

II. The probability of calling a false hypothesis true is $P(C|\bar{H})$.

This follows at once from the definition of C and $P(C|\bar{H})$.

Example:

To illustrate these ideas, suppose there is a certain disease which causes an elevation in the blood count, y, of some component. In particular, suppose the distribution of y among normal people has the probability density

$$p_n = 100e^{-100\ y}$$

and among those with disease

$$p_d = e^{-y}.$$

Let C be the set of all y greater than .01 and let H be the hypothesis that a certain person has the disease. Suppose that on the basis of a measurement of y the truth or falsity of H is to be decided according to the rule: H will be taken to be true if y is in C (i.e., $y \geq .01$) and false if y is not in C (i.e., $y < .01$). C is therefore the critical region.

$P(C|H)$ is given by

$$P(C|H) = \int_{.01}^{\infty} p_d(y)\, dy = \int_{.01}^{\infty} e^{-y}\, dy = e^{-.01} \approx .99.$$

and $P(C|\bar{H})$ by

$$P(C|\bar{H}) = \int_{.01}^{\infty} p_n(y)dy = \int_{.01}^{\infty} 100\, e^{-100y} dy = e^{-1}$$

since if the hypothesis were false the probability density of y would be $p_n(y)$.

The probability of a catastrophic error (in this case, calling the hypothesis false when true) is $1 - P(C|H)$ or .01. The probability of a false alarm is e-1. The possibility of keeping the catastrophic error probability so low while having a reasonable false alarm rate is due to the fact that $p_n(y)$ and $p_d(y)$ are so widely separated. If $p_n(y)$ were $1.1e^{-1.1y}$, the false alarm probability would be nearly unity and a measurement of y could not be used to distinguish between those with the disease and those without.

3.0 LUNAR SITE CERTIFICATION PROBLEM

The preceding ideas will now be extended to the lunar site certification problem.

Consider an area A on the lunar surface which is regarded as a potential site for a landing. Suppose some measurement* y of the properties of A is made and on the basis of this measurement a decision is to be made whether or not a landing is to be actually attempted in A. It is natural to impose the condition on any decision procedure that the probability of it leading to a catastrophic event under any possible circumstance should be very small (say less than .01).

A catastrophic event, in the present context, means a disaster overtaking the spacecraft on landing, while the phrase "any possible circumstance" refers to any possible surface which may be found on the moon.**

To put our condition in quantitative form, let f be the fraction of A which satisfies the requirements imposed on a landing point so that at any point of this fraction f a safe landing can be made; a conservative assumption will be made that an attempted landing at any point not meeting the requirements (i.e., outside of f) is disastrous or at least hazardous. Let

*"Measurement" here can mean the combination of a large number of single measurements.

**That is to say, any possible surface with respect to the specific properties measured. Of course, no statement can be made about the danger due to properties which are not observed. This qualification is to be understood throughout.

p(f) be then the probability that a landing within A may be safely made; for a random landing (i.e., if the landing point is chosen at random within A) p(f) is simply f; if there is some astronaut control, p(f) will generally be greater.

As before, we introduce a "critical region" C such that if the measurement y, of A lies within C an attempt will be made at a landing and otherwise, not. Finally let $P(C|f)$ be the probability that the measurement y would lie within C if f were the fraction of A meeting the requirements. Then our condition on C is this: if k is the maximum tolerable probability for a decision to lead to catastrophe, then we must have for <u>all</u> f

$$k \geq [1 - p(f)]\, P(C|f);$$

or, what is the same thing

$$k \geq \text{maximum } [1 - p(f)]\, P(C|f).$$

To see this, observe that for a decision to lead to a catastrophe two events must both occur:

(a) the measurement of the area must lie in the critical region C so that a landing attempt is made - probability $P(C|f)$

(b) the landing must be unsuccessful - probability $1 - p(f)$.

The joint probability of (a) and (b) is just $[1 - p(f)]\, P(C|f)$ and this must be no greater than k <u>for all</u> f. The reason for the phrase "for all f" is that since the actual f for any given area is not known one must allow for all possible cases.

For small f, $1 - p(f)$ may be near one, but $P(C|f)$ will be very small, i.e., one is unlikely to attempt a landing on such an area. For f large $P(C|f)$ may be near one, but $1 - p(f)$ will be very small; i.e., the landing is likely to be successful.

To define a false alarm probability, observe that a "false alarm" occurs whenever one rejects for landing an area on which a landing would have been attempted if the landable fraction f had been known. But those are the areas on which the probability of a catastrophic event is not greater than k; i.e., the areas whose f are such that $1 - p(f) \leq k$. Now, the probability of rejecting such areas is just $1 - P(C|f)$. Hence if r is the maximum false alarm probability then

$$r \leq [1 - P(C|f)]$$

where f is such that $|1 - p(f)| \leq k$.

Any requirement on the catastrophic error probability means that the critical region C must be sufficiently small. Decreasing the size of C, however, usually raises the false alarm probability. Thus a condition on both the catastrophic error and false alarm ultimately is a condition on the ability of the measurement to discriminate between a good or a bad landing area.

The following two examples will illustrate the above ideas:

1. Suppose that there is some property p of the lunar surface which can prove hazardous to a landing if its value exceeds some threshold. Let f be the fraction of the area A in which this threshold is not exceeded. Suppose n independent measurements of p are made, and that a landing attempt will be made only if in all n measurements threshold is not exceeded. If the maximum catastrophic error probability that can be tolerated is k, what is the smallest possible value of n? What is the corresponding false alarm rate? Assume random landing.

   A practical example is the measurement of bearing strength by means of accelerometer balls. Since in this case, the higher the bearing strength the better the surface, a natural condition is that one over the bearing strength be less than some threshold. (Another example, is the testing of a rocket for manned use. The criteria of acceptability is generally a prescribed number of consecutive successful firings, with f here being the probability of any given firing being successful.)

   By assumption, the critical region for landing is that all n measurements be less than threshold. Hence $P(C|f) = f^n$ and since $p(f) = f$, we have

   (1) $k \leq$ Maximum $(f^n)(1-f)$

   The maximum of $f^n(1-f)$ occurs at $f = \frac{n}{n+1}$, so that

   (2) $k \leq \left(\frac{n}{n+1}\right)^n \left(\frac{1}{n+1}\right)$

the least n satisfying the above inequality is for small k approximately

$$(3) \quad n = \frac{1}{ke}$$

using the fact that for large n, $e \approx (1 + \frac{1}{n})^{n+1}$.
From equation (2) one finds for k less than .05, it is sufficient to take n = 8, and for k = .01, it is sufficient to take n = 37.

The false alarm probability r is given by

$$r \leq \text{Max } (1 - f^n)$$

where f is such that $f > 1 - k$. In fact, a false alarm occurs whenever on a good surface (one for which $f \geq 1 - k$) a measurement greater than threshold is made.

But the probability of this is just $1 - f^n$.

Since Max $(1 - f^n)$ obviously occurs at $f = 1 - k$, we have

$$r \leq 1 - (1 - k)^n.$$

r is nearly constant for k sufficiently small ($\leq .05$).* From (2) we have $r \leq 1 - (1 - k)^{1/ke} = 1 - e^{-1/e} = .31$.

Actually the false alarm probability given above is somewhat pessimistic. For example if k = .05, it is true that in 31% of the cases an area for which f = .95 would be dismissed. But such a surface would a priori be regarded as barely tolerable. For a very good surface, say f = .98, the false alarm probability is $1 - (.98)^8$ or .145.

A decrease in false alarm probability can be obtained by widening the critical region so as to permit a landing even if one measurement is above threshold. However, for prescribed k, a greater number of measurements are required. In fact

$$k \leq \text{Max } (1 - f)\, P(C|f) = \text{Max } (1 - f)\, [f^n + n(1 - f) f^{n-1}]$$

*Assuming for each k, the minimum number of measurements are made.

A short calculation shows that it is sufficient to take

$$(4)\quad n = \frac{2}{ke} \text{ (approximately)},$$

just twice equation (2); and the resulting false alarm probability is approximately .17.

2. Suppose that it is known that some property of the lunar surface has a probability density $\theta e^{-\theta y}$ where the parameter θ is, however, unknown. Assume that if at any given point y is less than 1, then a safe landing may be made, but if $y > 1$, a landing is hazardous. A single measurement y_m of this property is made within A. If landing is random, find a critical region if the maximum acceptable probability of catastrophe is k. What is the corresponding false alarm rate?

It is clear that any sensible choice of critical regions (i.e., one which minimizes the false alarm rate while meeting the condition of the catastrophic error) must consist of all y_m such that $0 \leq y_m \leq y_o$ where y_o is some number that must be found. Indeed, if one is willing to accept the area for landing for a measurement value y_o one would naturally accept it for any better values (i.e., for any y_m less than y_o). To find y_o, note that if f is the fraction of A for which $y < 1$ then

$$f = \theta \int_0^1 e^{-\theta y}\, dy = 1 - e^{-\theta}.$$

Hence, since $C = \{\text{all y such that } 0 \leq y \leq y_o\}$

$$P(C|f) = \theta \int_0^{y_o} e^{-\theta y}\, dy = 1 - e^{-\theta y_o};$$

and the probability of catastrophic error is

$$(1)\quad e^{-\theta}(1 - e^{-\theta y_o}),$$

since $1-p(f) = 1 - f = 1 - (1 - e^{-\theta}) = e^{-\theta}$.

The quantity (1) must be maximized over θ, and y_o chosen so that this maximum is less than k. The maximum occurs when

$$e^{-\theta y_o} = \frac{1}{1 + y_o}$$

Substituting in (1), it is readily shown that k, the maximum catastrophic error probability, is related to y_o by $y_o = \frac{ek}{1 - ke}$ (very nearly).

In particular for $k = .05$, $y_o = .157$, and $k = .01$, $y_o = .0276$. The false alarm probability for $k = .05$ is

$$r < \theta \int_{y_o}^{\infty} e^{-y\theta}\, dy = e^{-y_o\theta}$$

where $y_o = .157$ and θ is such that $1 - e^{-\theta} = .95$. Thus $r < .624$. Similarly for $k = .01$, we find that $r < .88$.

These high false alarm probabilities are a consequence of using only a single measurement of the area as a basis for our statistical decision procedure. They may be reduced to an arbitrarily low level by making a sufficient number of measurements.

SUMMARY AND CONCLUSIONS

The basic ideas of statistical decision theory have been discussed and an extension has been made to the problem of lunar site certification. A statistical procedure has been presented which guarantees that the probability that a decision will lead to disaster is less than any prescribed value; this result is independent of the actual characteristics of the surface, a very desirable thing since these characteristics are not known.

It was pointed out, however, that lowering the catastrophic error probability increases the false alarm rate. To keep both in reasonable limits imposes a requirement on how extensive measurements of the lunar surface must be.

5.0 ACKNOWLEDGEMENTS

Thanks are due to Dennis James and Charles Byrne for many helpful discussions.

This material will, in modified form, appear also as part of a report "Statistical Analysis of Site Survey by the Unmanned Program," written jointly with C. J. Byrne, R. F. Fudali, and T. L. Powers.

1112-MAL-WMT